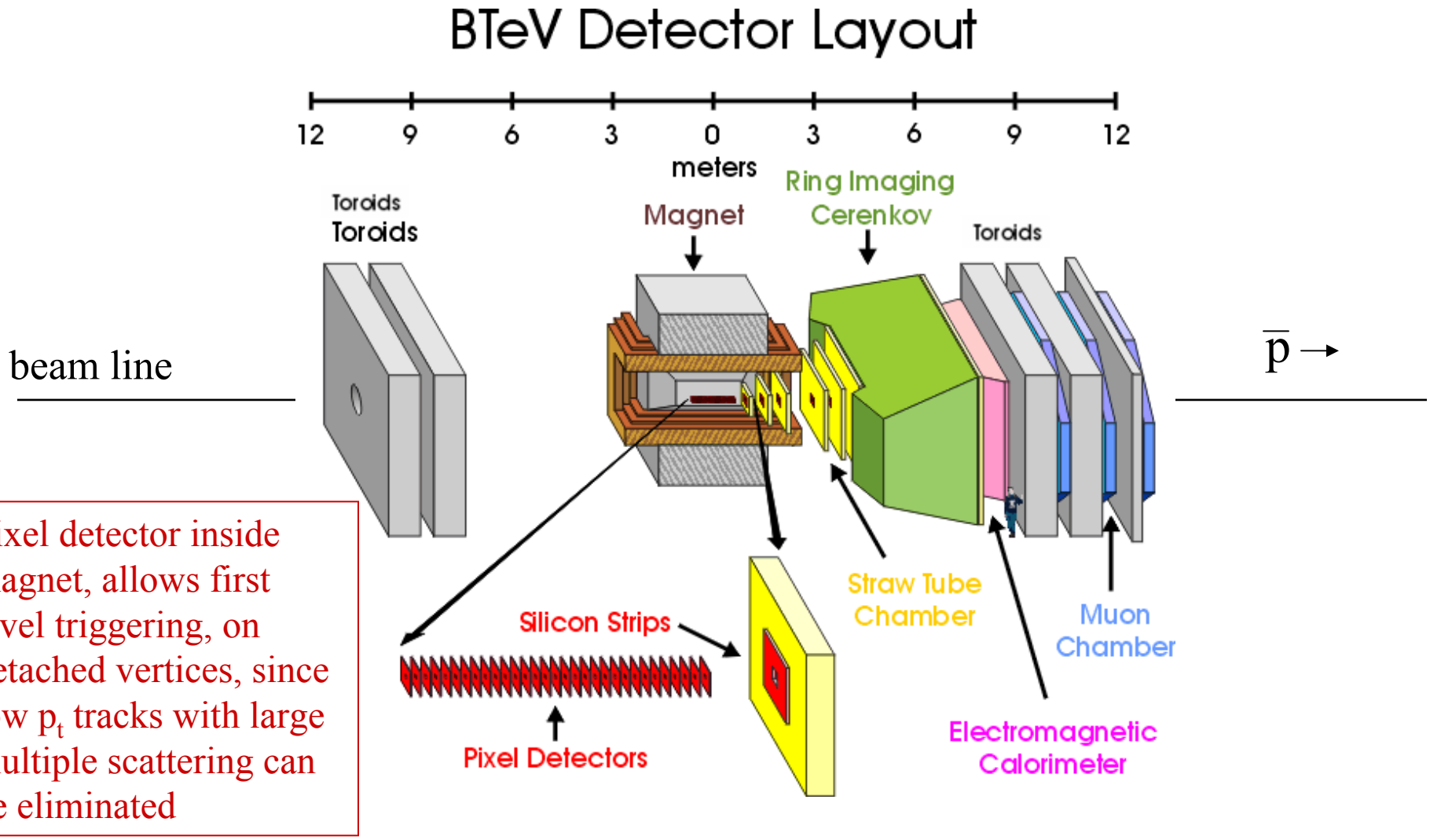
BTeV Trigger and DAQ Innovations

Margaret Votava
Fermilab
On Behalf of BTeV DAQ and Trigger Groups

The BTeV Detector



Pixel detector inside magnet, allows first level triggering, on detached vertices, since low p_t tracks with large multiple scattering can be eliminated

BTeV Background

- B-physics experiment proposed at Fermilab
  - www-btev.fnal.gov
  - 170 physicists from 30 universities/institutions worldwide (and growing)
  - $193M construction project ($141M base + $52M contingency)
    - DAQ and Trigger were 2 separate subprojects accounting for 17% of the project cost combined.
- Timeline
  - In R&D since 1996 (!)
  - Construction to have begun in FY2005
  - Two stage delivery (to match funding profile) with ~50% capacity in FY09 and the remaining in FY10.
  - Successfully passed CD2/CD3a (ie, limited construction funds available) review in December 2004.
  - Cancelled abruptly in February, 2005 by the Department of Energy

Online Requirements

- The challenge for the BTeV trigger and data acquisition system is to reconstruct particle tracks and interaction vertices for **EVERY** interaction that occurs in the BTeV detector, and to select interactions with *B* decays.
- The trigger performs this task using 3 levels, referred to as Levels 1, 2, and 3:
  "L1" – looks at every interaction and rejects at least 98% of min. bias background
  "L2" – uses L1 computed results & performs more refined analyses for data selection
  "L3" – rejects additional background and performs data-quality monitoring

Reject > 99.9% of background. Keep > 50% of *B* events.

- The data acquisition system saves all of the data in memory for as long as necessary to analyze each interaction, and moves data to L2/3 processing units and archival data storage for selected interactions.
  - Complex algorithms => long latencies (1 msec for L1)
- The key ingredients that make it possible to meet this challenge:
  - BTeV pixel detector with its exceptional pattern recognition capabilities
  - Rapid development in technology – FPGAs, processors, networking
- I/O rates
  - 21M channels, dominated by Pixel detector
  - Input rate 2.5MHz x 200KBytes/event = 500 Gbytes/sec.
  - Output rate = 2.5 KHz x 80 Kbytes/event = 200 Kbytes/sec
    - 1 PByte/year

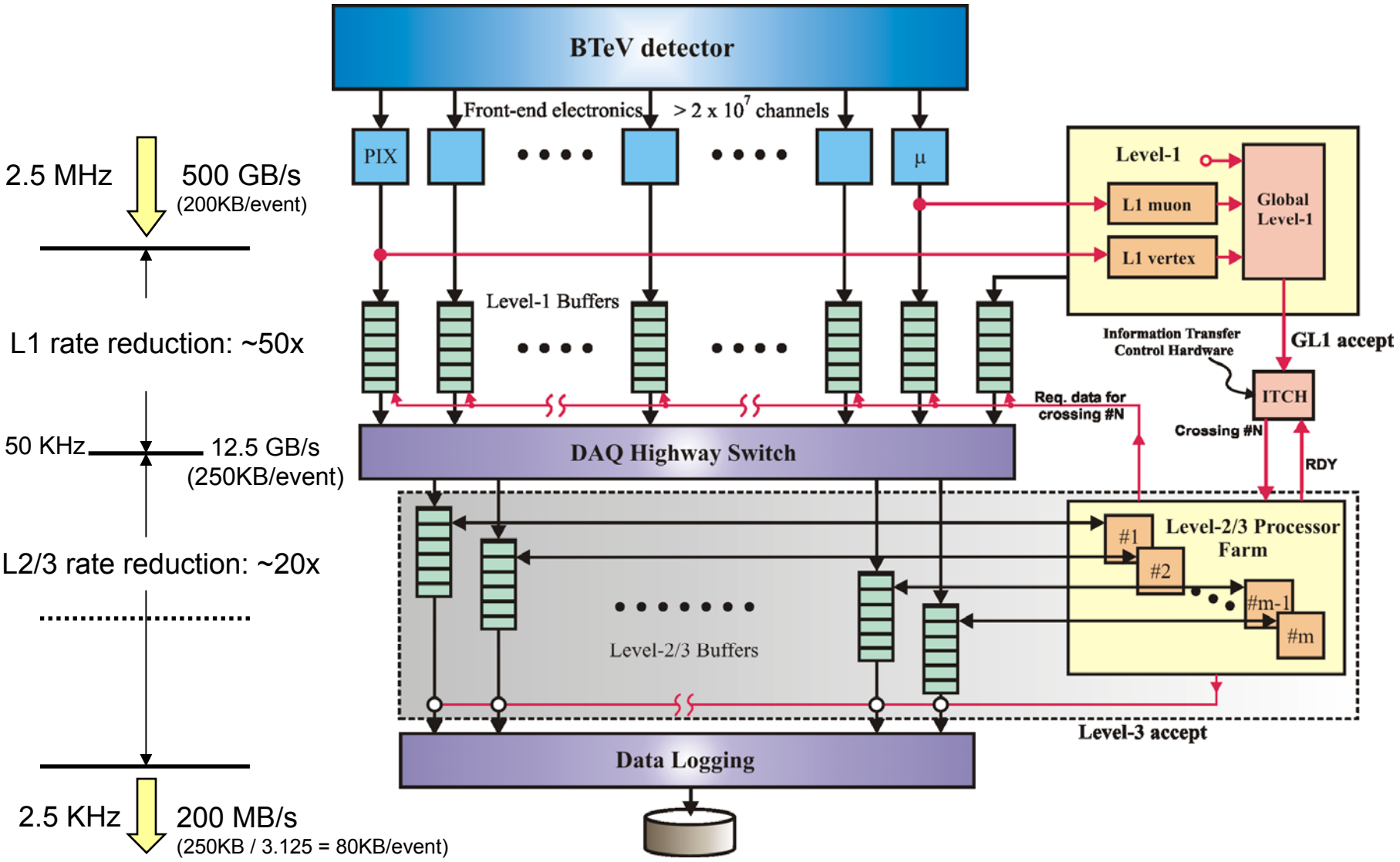
How to do this Affordably?

- Fast data links between detector and online system
- Sophisticated L1 track findingTrigger
- Commodity hardware wherever possible.
- Use inexpensive DRAM-type buffer memory off detector
- Only point-point serial links (copper and optical)
- Subdivide system into 8 parallel highways

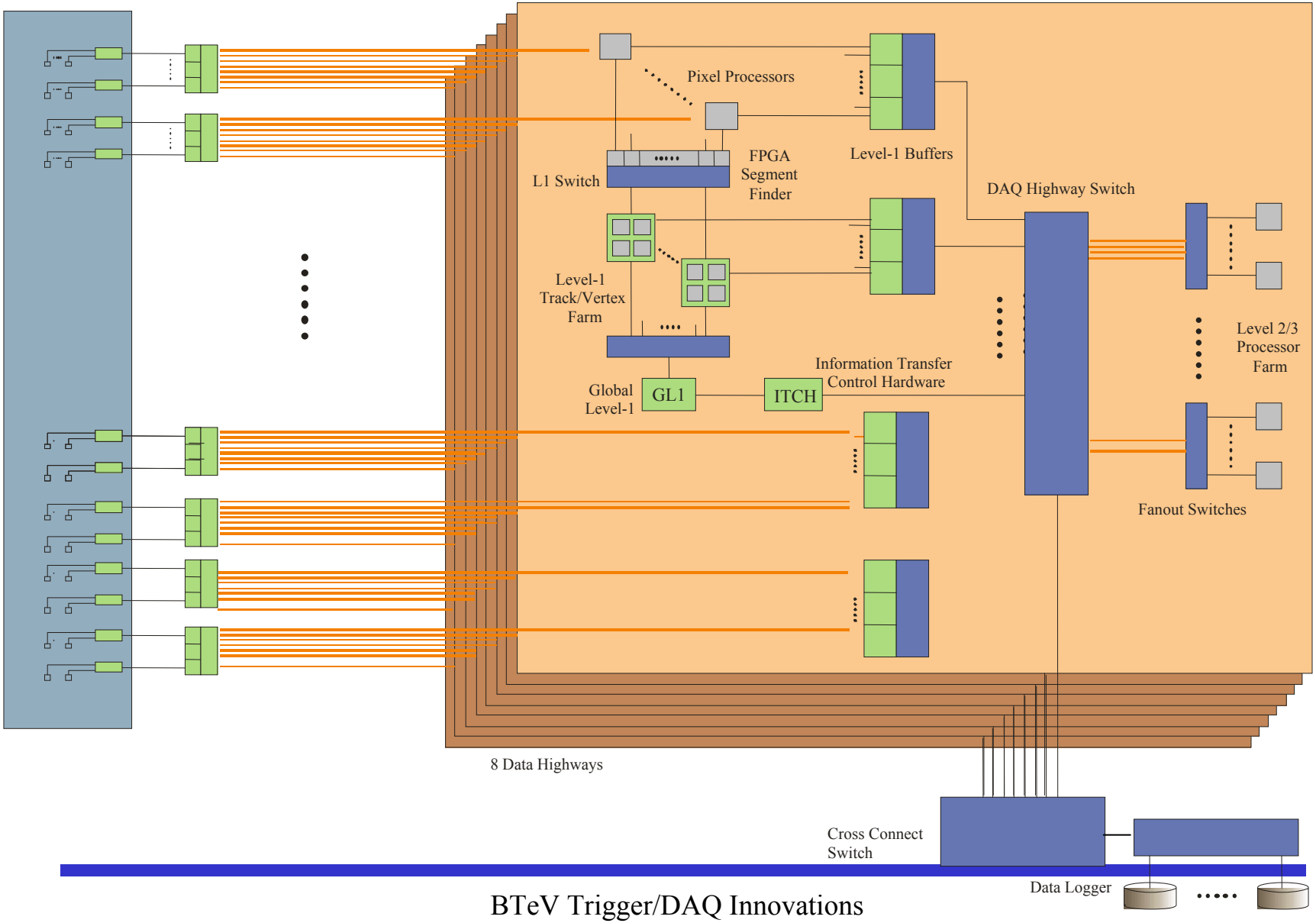
Online Architecture Credible

- Detailed work break down structure for complete experiment.
  - Fully burdened resource loaded schedule
  - ~6000 activities for trigger/DAQ.
  - Risk analysis understood and mitigation strategies planned.
  - Included both cost and schedule contingency. Online system completed at least 9 months before it was needed.
- Heavily reviewed.
  - Went before P5 subpanel (twice!)
  - External CD0, CD1, and CD2/3a, each preceeded by an internal director's review.
  - BTeV passed the final review – CD2/3a with notably favorable comments from the review committee.
- Prototyping
  - L1 Trigger prototyped
  - Prepilot being assembled

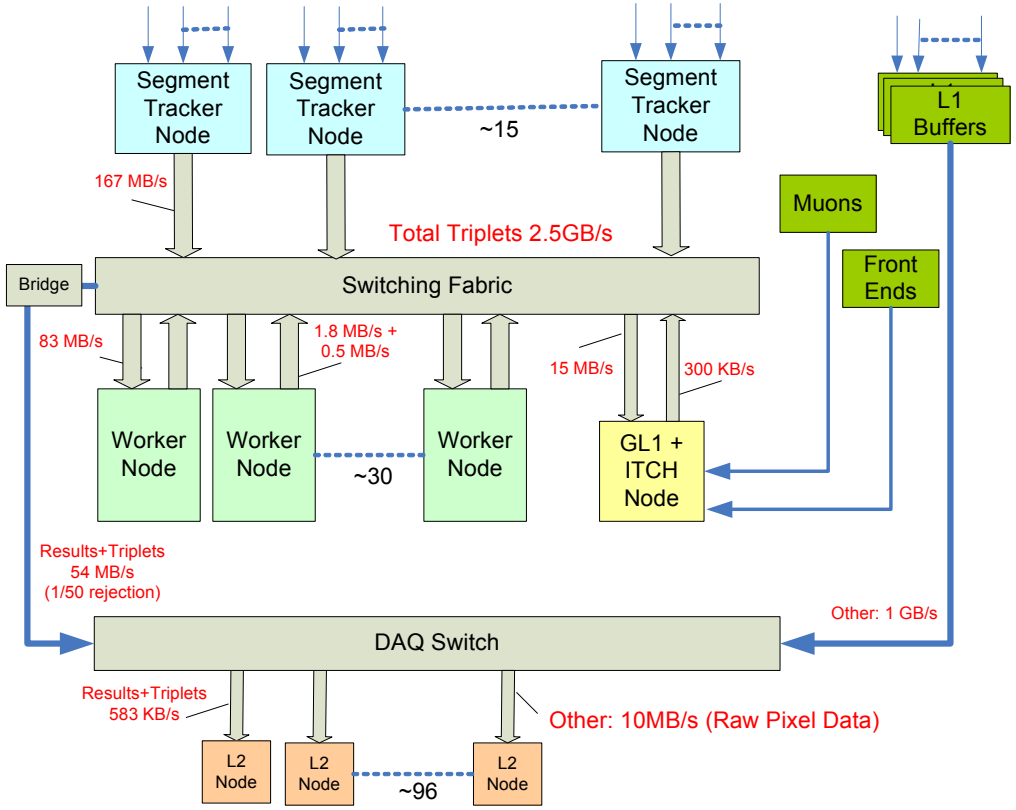
Block Diagram of Trigger & DAQ Data Flow



Online Architecture



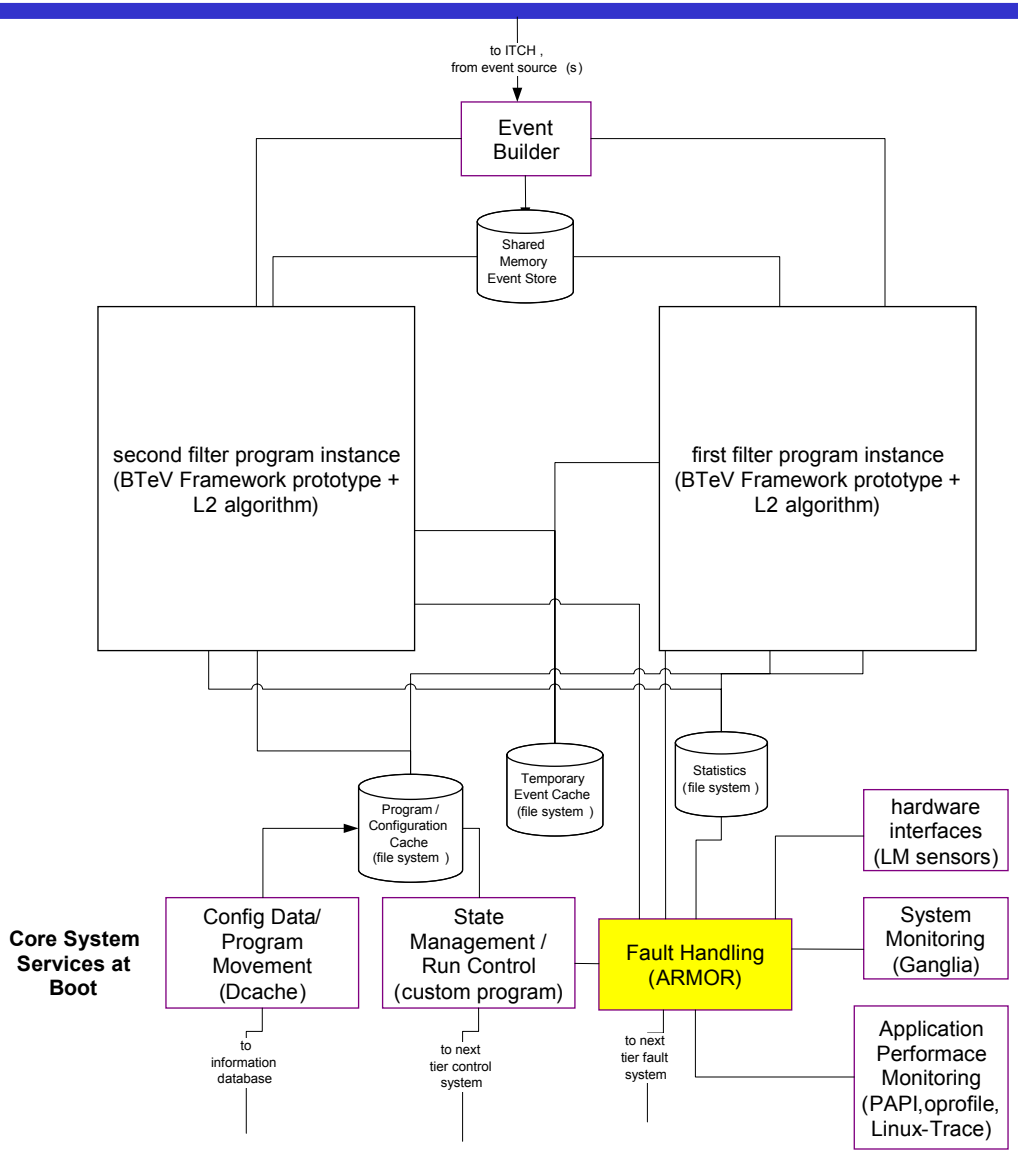
L1 Trigger


Segment Tracker Node
Segment Tracker Node
~15
Segment Tracker Node
L1 Buffers
167 MB/s
Total Triplets 2.5GB/s
Muons
Front Ends
Bridge
Switching Fabric
83 MB/s
1.8 MB/s + 0.5 MB/s
15 MB/s
300 KB/s
Worker Node
Worker Node
~30
Worker Node
GL1 + ITCH Node
Results+Triplets 54 MB/s (1/50 rejection)
Other: 1 GB/s
DAQ Switch
Results+Triplets 583 KB/s
Other: 10MB/s (Raw Pixel Data)
L2 Node
L2 Node
~96
L2 Node

RTES

- Talked about before
  - Fault tolerant adaptive embedded systems
  - Montreal talk by L. Piccoli – using Armors (one component) for process management
  - Collaboration of physicists and computer scientists
  - Near end of year 4 of 5 year grant
- Talked about again
  - Mike Haney (UofI) – BTeV and Beyond

Worker Node Software Architecture



Partitioning

- Defined as ability to run multiple data acquisition sessions in parallel.
- Commissioning goals
  - Test out subdetectors and electronics standalone
  - Test out subdetectors and electronics in gangs
  - Data collection could be a few sparse events all the way to full rate.
- Operational goals
  - Obey prime directive => collect maximum amount of physics events.
  - Corral spare cycles on the online trigger farm for offline processing in a non-disruptive way.
  - Test out new trigger algorithms
- Logical concept
  - Myriad of implementation possibilities given the 8 highway architecture and run time configurable routing tables in the DCBs

Paritioning Proposal

- Still being debated in collaboration
- Human run coordinator
  - Establish number of working highways
  - Coordinate data taking runs during this period (e.g., knowledge of who will need pixel trigger)
  - Start trigger
- Commissioning
  - Each Data taker
    - Select data sources (in units of L1 buffers) for read/write or read only (can't reset)
    - Select data syncs (in units of regional managers)
    - Select set of trigger tables
- Operations
  - Worker nodes can be manually assigned to offline partition (for long down times) in units of regional managers
  - Worker nodes can automatically shift to offline partition as luminosity decreases (can be detected by RTES)

Highlights

- Credible Design
- Standardizing inputs to DAQ/Trigger as early as possible. DCBs were single entry point, developed by a centralized institution.
- Highway architecture
  - Individual highways operating at 1/8 of full rate
  - Control overhead more manageable
  - Network bandwidth used for efficiently (larger packet sizes)
- Asynchronous L1 vertex tracker on COMMODITY hardware.
  - lower cost
  - lower risk
  - requires an increase in power & cooling
  - easier to build (less custom hardware to design)
  - requires less labor (less engineering & no DSP programming)
  - easier upgrade path (easier to add or replace processors)
- Reliable, fault adaptive system .
- Offline capabilities on online farm.

Contacts

- Contact us:
  - Btev-daq@fnal.gov
  - Btev_trigger@fnal.gov
  - votava@fnal.gov
- Talk to us
  - Mike Haney
    - The RTES Project - BTeV, and Beyond (S9-1)
  - Mike Wang
    - A Commodity Solution Based High Data Rate Asynchronous Trigger System for Hadron Collider (S14-1)
  - Jin Yuan Wu
    - The Application of Tiny Triplet Finder (TTF) in BTeV Pixel Trigger (S6-3)
    - Integrated Upstream Parasitic Event Building Architecture for BTeV Level 1 Pixel Trigger System (P8-3)
  - Luciano Piccoli
    - Unrelated talk
- Surf the web
  - http://www-btev.fnal.gov – BTeV home page include links to TDR
  - http://www-btev.fnal.gov/public/hep/detector/rtes/ - RTES home page